GENERIC CODING AND ITS USE IN MANUSCRIPT PREPARATION

by Marianne Brogan
American Chemical Society
Columbus, Ohio

Abstract: The role of generic coding in manuscript preparation and production is discussed.

In the past 10 years the changes in production of scholarly publications-- indeed all publications--have been dramatic: from traditional hot metal to photocomposition, from conventional keyboarding to increasing use of technology, composition houses have completely reversed their images, adapting their procedures to the use of high technology. But the changes they have to make are not over: they are being besieged by customers to accept input via word processing interfaces. Composition houses are accepting the challenge, so much so that at an April 1983 meeting of the National Composition Association it was recognized that within 2 years ca. 90% of their work would be delivered by using electronic methods.

During the past 5 years, the same pressures have been exerted on publishers: mild at first, increasing in strength, until now on all sides the question is asked by authors: Why can't you use my word-processing output?

Thus, publishers are in the middle, being the customer but not the producer. When manuscripts have been prepared by use of electronic media, direct use by the publisher of the initial keystrokes should result in significant savings, among these being savings in proofreading time on the part of the author and reduction of duplication of effort and proofreading expense on the part of the publisher. However, these benefits will accrue only if the paper is prepared in a form suitable for easy manipulation by the publisher. How do we bridge the gap, to get one community in a position to serve the other but in such a way that it itself is well served? There are answers to these questions, but the answers are not easy and are not uniform. Coding plays an important part.

The American Chemical Society has been involved with generic coding since the 1960s. In 1967, computerization of indexing functions at Chemical Abstracts Service (CAS) demanded identification codes. In the first 7 years of use, the application was directed almost exclusively toward data base management. However, by the time the decision was made to incorporate primary journal production into the CA computer environment, additional codes to support such an application had been created.

1975 marked the production of our first journal on the system designed for us by CAS staff. The system was modified over the years, and enhancements continue, but the software that supports ACS primary journals now, and supports data base creation and implementation, is essentially the same that was introduced in 1975--and generic codes were as essential part of that system.

The important features were threefold:

(1) Data were identified and input in the most elemental form.
(2) Data were kept separate from format. requirements.
(3) Processes were automated to the greatest degree practical,rwith flexibility and expandability an integral part.

As a result, the data base created could be used for multiple purposes. Those purposes currently consist of production of our primary journals in traditional form, including automatic table of contents and author index generation, as well as creation of a data base of published material--a full text file that can be used for full-text searching and document delivery. That text file has been privately available for several years and is being made publicly available in June.

Generic coding permits multiple formats and multiple uses. Of the 18 journals published by ACS using this system, it is possible to do a single keyboarding and have that data output in at least 41 different formats: that of the three printers to which our word-processing system is linked, Xerox 9700 output of the 18 different journal formats, APS-5 (our phototypesetter) output of the 18 different journal formats, computer line printer, and that of our electronic data base.

All of which is interesting, but it does not resolve the issue at hand: taking authors' manuscripts in electronic form, editing them, and using the electronic data to drive the composition system. We have in fact done so since 1974, and with automatic pagination on our present system

Infrared Spectrometry

Robert S. McDonald

General Electric Corporate Research and Development Center,

INTRODUCTION

This review covers mainly publications cited in *Chemical Abstracts*, volumes 92–95 (1980–81), through the December 28, 1981 issue (Jan. 11, 1982 issue of *CA Selects*).

Selection of References. The initial selection was based on *CA Selects: Infrared Spectrometry (Physicochemical Aspects)* and *Infrared Spectroscopy (Organic Aspects)*. *CA Selects* have proved to be a far more useful source for this review than *CA Search* which was used for the previous five reviews. The main advantage of *CA Selects* is that they contain complete abstracts rather than citations and keywords only. A disadvantage has been that *CA Selects* were not available to this reviewer in computer readable form like *CA Search*. However, most of the citations have been taken from the SDC *CA* database in computer readable form (*1*). The citations from SDC were parsed and rearranged into the standard citation form for *Analytical Chemistry* by computer. Thus, in a sense most of the bibliography of this review has been effectively keyed by *Chemical Abstracts* as has been the case for the three previous reviews.

This year, the review is again being submitted to *Analytical Chemistry* in computer readable form. A new version of the computer program *LISE (Literature Search and Edit)* which has been described in previous reviews was used to manipulate the citations.

Figure 1. Sample from a journal article produced by use of author-supplied magnetic tape. (Reproduced with the permission of the author from Analytical Chemistry. Copyright 1982, American Chemical Society.)

since 1978, but under one condition: the author meet our specifications. And indeed our specifications have been very exacting: at present, 9 track, 1600 bpi tape, ASCII character representation, generic coding, and character set coding (Figure 1 shows a paper produced by using the protocol established).

This brings up the often repeated question--what level editing?--but in a different context: what level coding? Do we as editors and publishers want our authors to supply us with manuscripts with all codes inserted? Why not, if it is done properly? There will always be manuscripts with coding demands so exacting that editors would prefer a hands-off approach, but with use of function keys with programmed patterns, a moderate level of generic coding on the part of the author is not outside of the realm of possibility.

The American National Standards Institute (ANSI) has almost completed a standard on the subject of generic coding. The Association of American Publishers (AAP) has instituted a project for the "Development of Publishing Industry Standards and Author Guidelines on Electronic Manuscript Preparation". Both efforts should combine in the next few years to provide both authors and publishers a road map, one that will help the participants in the communications dialogue follow the communications routes that are now available to us.